SORT 40 (1) January-June 2016, 55-88

A test for normality based on the empirical distribution function

Hamzeh Torabi[1], Narges H. Montazeri[1] and Aurea Grané[2]

Abstract

In this paper, a goodness-of-fit test for normality based on the comparison of the theoretical and empirical distributions is proposed. Critical values are obtained via Monte Carlo for several sample sizes and different significance levels. We study and compare the power of forty selected normality tests for a wide collection of alternative distributions. The new proposal is compared to some traditional test statistics, such as Kolmogorov-Smirnov, Kuiper, Cramér-von Mises, Anderson-Darling, Pearson Chi-square, Shapiro-Wilk, Shapiro-Francia, Jarque-Bera, SJ, Robust Jarque-Bera, and also to entropy-based test statistics. From the simulation study results it is concluded that the best performance against asymmetric alternatives with support on the whole real line and alternative distributions with support on the positive real line is achieved by the new test. Other findings derived from the simulation study are that SJ and Robust Jarque-Bera tests are the most powerful ones for symmetric alternatives with support on the whole real line, whereas entropy-based tests are preferable for alternatives with support on the unit interval.

MSC: 62F03, 62F10.

Keywords: Empirical distribution function, entropy estimator, goodness-of-fit tests, Monte Carlo simulation, Robust Jarque-Bera test, Shapiro-Francia test, SJ test; test for normality.

1. Introduction

Let $X_1,\ldots,X_n$ be a n independent an identically distributed (iid) random variables with continuous cumulative distribution function (cdf) $F(.)$ and probability density function (pdf) $f(.)$. All along the paper, we will denote the order statistic by $(X_{(1)},\ldots,X_{(n)})$. Based on the observed sample $x_1,\ldots,x_n$, we are interested in the following goodness-of-fit test for a location-scale family:

[1] Statistics Department, Yazd University, 89175-741, Yazd, Iran, htorabi@yazd.ac.ir, nmontazeri@stu.yazd.ac.ir

[2] Statistics Department, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe, Spain, aurea.grane@uc3m.es

Received: April 2015

Accepted: February 2016

$$\begin{cases} H_0: \ F \in \mathscr{F} \\ H_1: \ F \notin \mathscr{F} \end{cases} \tag{1}$$

where $\mathscr{F} = \left\{F_0(.;\boldsymbol{\theta}) = F_0\left(\frac{x-\mu}{\sigma}\right) \mid \boldsymbol{\theta} = (\mu,\sigma) \in \Theta\right\}$, $\Theta = \mathbb{R} \times (0,\infty)$ and μ and σ are unspecified. The family $\mathscr{F}$ is called location-scale family, where $F_0(.)$ is the standard case for $F_0(.;\boldsymbol{\theta})$ for $\boldsymbol{\theta} = (0,1)$. Suppose that $f_0(x;\boldsymbol{\theta}) = \frac{1}{\sigma} f_0\left(\frac{x-\mu}{\sigma}\right)$ is the corresponding pdf of $F_0(x;\boldsymbol{\theta})$.

The goodness-of-fit test problem for location-scale family described in (1) has been discussed by many authors. For instance, Zhao and Xu (2014) considered a random distance between the sample order statistic and the quasi sample order statistic derived from the null distribution as a measure of discrepancy. On the other hand, Alizadeh and Arghami (2012) used a test based on the minimum Kullback-Leibler distance. The Kullback-Leibler divergence measure is a special case of a ϕ-divergence measure (2) for $\phi(x) = x\log(x) - x + 1$ (see p. 5 of Pardo, 2006 for details). Also ϕ-divergence is a special case of the ϕ-disparity measure. The ϕ-disparity measure between two pdf's f_0 and f is defined by

$$D_\phi(f_0, f) = \int \phi\left(\frac{f_0(x;\boldsymbol{\theta})}{f(x)}\right) f(x)\, dx, \tag{2}$$

where $\phi : (0,\infty) \to [0,\infty)$ is assumed to be continuous, decreasing on $(0,1)$ and increasing on $(1,\infty)$, with $\phi(1) = 0$ (see p. 29 of Pardo, 2006 for details). In ϕ-divergence, ϕ is a convex function.

Inspired by this idea, in this paper we propose a goodness-of-fit statistic to test (1) by considering a new proximity measure between two continuous cdf's. The organization of the paper is as follows. In Section 2 we define the new measure H_n and study its properties as a goodness-of-fit statistic. In Section 3 we propose a normality test based on H_n and find its critical values for several sample sizes and different significance levels. In Section 4 we review forty normality tests, including the most traditional ones such as Kolmogorov-Smirnov, Cramér-von Mises, Anderson-Darling, Shapiro-Wilk, Shapiro-Francia, Pearson Chi-square, among others, and in Section 5 we compare their performances to that of our proposal through a wide set of alternative distributions. We also provide an application example where the Kolmogorov-Smirnov test fails to detect the non normality of the sample.

2. A new discrepancy measure

In this section we define a discrepancy measure between two continuous cdf's and study its properties as a goodness-of-fit statistic.

Definition 2.1 *Let X and Y be two absolutely continuous random variables with cdf's F_0 and F, respectively. We define*

$$D(F_0,F) = \int_{-\infty}^{\infty} h\left(\frac{1+F_0(x;\boldsymbol{\theta})}{1+F(x)}\right) dF(x) = E_F\left[h\left(\frac{1+F_0(X;\boldsymbol{\theta})}{1+F(X)}\right)\right], \tag{3}$$

where $E_F[.]$ is the expectation under F and $h:(0,\infty)\to\mathbb{R}^+$ is assumed to be continuous, decreasing on $(0,1)$ and increasing on $(1,\infty)$ with an absolute minimum at $x=1$ such that $h(1)=0$.

Lemma 2.2 *$D(F_0,F)\geq 0$ and equality holds if and only if $F_0=F$, almost everywhere.*

Proof. Using the non-negativity of function h, we have $D(F_0,F)\geq 0$. It is clear that $F_0=F$ implies $D(F_0,F)=0$. Conversely, if $D(F_0,F)=0$, since h has an absolute minimum at $x=1$, then $F_0=F$. ∎

Let us return to the goodness-of-fit test problem for a location-scale family described in (1). Firstly, we estimate μ and σ by their maximum likelihood estimators (MLEs), i.e., $\hat{\mu}$ and $\hat{\sigma}$, respectively, and we take $z_i=(x_i-\hat{\mu})/\hat{\sigma}$, $i=1,\ldots,n$. Note that in this family, $F_0(x_i;\hat{\mu},\hat{\sigma})=F_0(z_i)$. Secondly, consider the empirical distribution function (EDF) based on data x_i, that is

$$F_n(t)=\frac{1}{n}\sum_{j=1}^{n}\mathrm{I}_{[x_j\leq t]},$$

where I_A denotes the indicator of an event A. Then, our proposal is based on the ratio of the standard cdf under H_0 and the EDF based on the x_i's. Using (3) with $F=F_n$, $D(F_0,F_n)$ can be written as

$$\begin{aligned}
\mathrm{H}_n := D(F_0,F_n) &= \int_{-\infty}^{\infty} h\left(\frac{1+F_0(x;\hat{\mu},\hat{\sigma})}{1+F_n(x)}\right) dF_n(x)\\
&= \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(x_{(i)};\hat{\mu},\hat{\sigma})}{1+F_n(x_{(i)})}\right)\\
&= \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(z_{(i)})}{1+i/n}\right)
\end{aligned}$$

Under H_0, we expect that $F_0(t;\hat{\mu},\hat{\sigma})\approx F_n(t)$, for every $t\in\mathbb{R}$ and $1+F_0(t;\hat{\mu},\hat{\sigma})\approx 1+F_n(t)$. Note that, since $h(1)=0$, we expect that $h\big((1+F_0(t))/(1+F_n(t))\big)\approx 0$ and thus H_n will take values close to zero when H_0 is true. Therefore, it seems justifiable that H_0 must be rejected for large values of H_n. Some standard choices for h are: $h(x) = (x-1)^2/(x+1)^2, x\log(x)-x+1, (x-1)\log(x), |x-1|$ or $(x-1)^2$ (for more examples, see p. 6 of Pardo, 2006 for details).

Proposition 2.3 *The support of* H_n *is* $[0, \max(h(1/2), h(2))]$.

Proof. Since $F_0(.)$ and F_n are cdf's and take values in $[0,1]$, we have that

$$1/2 \leq \frac{1+F_0(y)}{1+F_n(y)} \leq 2, \quad y \in \mathbb{R}.$$

Thus

$$0 \leq h\left(\frac{1+F_0(y)}{1+F_n(y)}\right) \leq \max(h(1/2), h(2))$$

Finally, since $\mathrm{H_n}$ is the mean of $h(.)$ over the transformed data, the result is obtained. ■

Proposition 2.4 *The test statistic based on* H_n *is invariant under location-scale transformations.*

Proof. The location-scale family is invariant under the location-scale transformations of the form $g_{c,r}(X_1,\ldots,X_n) = (rX_1+c,\ldots,rX_n+c)$, $c \in \mathbb{R},\ r>0$, which induces similar transformations on $\Theta : g_{c,r}(\boldsymbol{\theta}) = (r\mu+c, r\sigma)$ (See Shao, 2003). The estimator $T_0(X_1,\ldots,X_n)$ for μ is location-scale invariant if

$$T_0(rX_1+c,\ldots,rX_n+c) = rT_0(X_1,\ldots,X_n)+c, \quad \forall r>0, c \in \mathbb{R},$$

and the estimator $T_1(X_1,\ldots,X_n)$ for σ is location-scale invariant if

$$T_1(rX_1+c,\ldots,rX_n+c) = rT_1(X_1,\ldots,X_n), \quad \forall r>0, c \in \mathbb{R}.$$

We know that MLE of μ and σ are location-scale invariant for μ and σ, respectively. Therefore under H_0, the distribution of $Z_i = (X_i - \hat{\mu})/\hat{\sigma}$ does not depend on μ and σ.

If G_n is the EDF based on data z_i, then

$$G_n(z_i) = \frac{1}{n}\sum_{j=1}^{n} \mathrm{I}_{[z_j \leq z_i]} = \frac{1}{n}\sum_{j=1}^{n} \mathrm{I}_{[x_j \leq x_i]} = F_n(x_i),$$

therefore

$$\mathrm{H}_n = \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(x_{(i)};\hat{\mu},\hat{\sigma})}{1+F_n(x_{(i)})}\right) = \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(z_{(i)})}{1+G_n(z_{(i)})}\right).$$

Since the statistic H_n is a function of z_i, $i = 1,\ldots,n$, is location-scale invariant. As a consequence, the null distribution of H_n does not depend on the parameters μ and σ. ■

Proposition 2.5 *Let F_1 be an arbitrary continuous cdf in H_1. Then under the assumption that the observed sample have cdf F_1, the test based on H_n is consistent.*

Proof. Based on Glivenko-Cantelli theorem, for n large enough, we have that $F_n(x) \simeq F_1(x)$, for all $x \in \mathbb{R}$. Also $\hat{\mu}$ and $\hat{\sigma}$ are MLEs of μ and σ, respectively, and hence are consistent. Therefore

$$\begin{aligned}
\mathrm{H}_n &= \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(x_{(i)};\hat{\mu},\hat{\sigma})}{1+F_n(x_{(i)})}\right) = \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(x_i;\hat{\mu},\hat{\sigma})}{1+F_n(x_i)}\right) \\
&\simeq \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(x_i;\hat{\mu},\hat{\sigma})}{1+F_1(x_i)}\right) \simeq \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(x_i,\mu,\sigma)}{1+F_1(x_i)}\right) \\
&\to E_{F_1}\left[h\left(\frac{1+F_0(X,\mu,\sigma)}{1+F_1(X)}\right)\right] =: D(F_0,F_1), \text{ as } n\to\infty,
\end{aligned}$$

where $E_{F_1}[.]$ is the expectation under F_1, and μ and σ^2 are, respectively, the expectation and variance of F_1. Note that the convergence holds by the law of large numbers and $D(F_0,F_1)$ is a divergence between F_0 and F_1. So the test based on H_n is consistent. ■

3. A normality test based on H_n

Many statistical procedures are based on the assumption that the observed data are normally distributed. Consequently, a variety of tests have been developed to check the validity of this assumption. In this section, we propose a new normality test based on H_n.

Consider again the goodness-of-fit testing problem described in (1), where now $f_0(x;\mu,\sigma) = 1/\sqrt{2\pi\sigma^2}e^{-(x-\mu)^2/2\sigma^2}$, $x \in \mathbb{R}$, in which $\mu \in \mathbb{R}$ and $\sigma > 0$ are both unknown, and $F_0(.;\mu,\sigma)$ is the corresponding cdf, where $F_0(.)$ is the standard case for $F_0(.;0,1)$.

First we estimate μ and σ by their maximum likelihood estimators (MLEs), i.e., $\hat{\mu} = \bar{x} = 1/n\sum_{i=1}^{n} x_i$ and $\hat{\sigma}^2 = s^2 = 1/(n-1)\sum_{i=1}^{n}(x_i-\bar{x})^2$, respectively. Let $z_i = (x_i-\bar{x})/s$, $i = 1,\ldots,n$. Then, the test statistic for normality is:

$$\mathrm{H}_n = \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(x_{(i)},\bar{x},s)}{1+F_n(x_{(i)})}\right) = \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(z_{(i)})}{1+i/n}\right), \tag{4}$$

where

$$h(x) = \left(\frac{x-1}{x+1}\right)^2. \tag{5}$$

Note that $h : (0,\infty) \to \mathbb{R}^+$ is decreasing on $(0,1)$ and increasing on $(1,\infty)$ with an absolute minimum at $x = 1$ such that $h(1) = 0$ (see Figure 1). We selected this function h, because based on simulation study, it is more powerful than other functions h. For example, we considered $h_2(x) := x\log(x) - x + 1$ for comparison with $h_1(x) := \left(\frac{x-1}{x+1}\right)^2$ (see Tables 6 and 7).

Corollary 3.1 *The support of* H_n *is* $[0, 0.11]$.

Proof. From Proposition 2.3 and Figure 1, $\max(h(1/2), h(2)) = 0.11$. ∎

Table 1 contains the upper critical values of H_n, which have obtained by Monte Carlo from 100000 simulated samples for different sample sizes n and significance levels $\alpha = 0.01, 0.05, 0.1$.

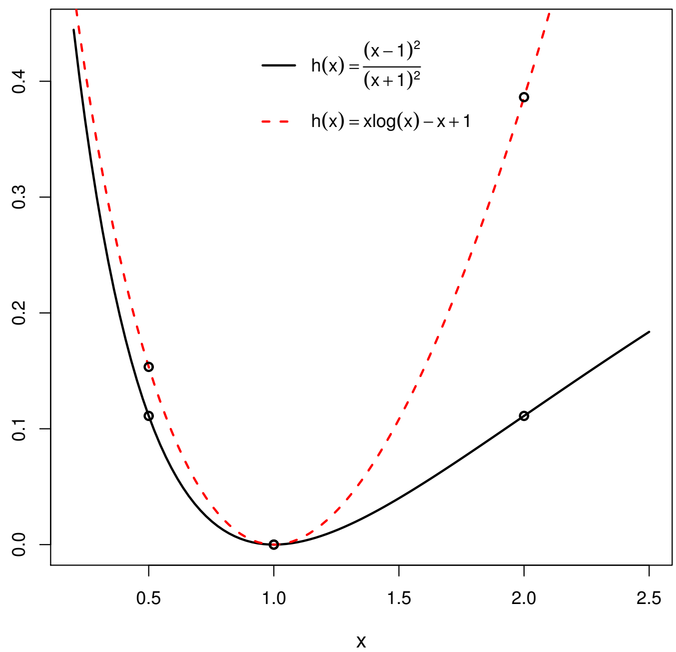

Figure 1: *Plot of function h.*

Table 1: *Critical values of* H_n *for* $\alpha = 0.01, 0.05, 0.1$.

n / α	5	6	7	8	9	10	15	20	25	30	40	50
0.01	.0039	.0035	.0030	.0026	.0023	.0021	.0014	.0011	.0008	.0007	.0005	.0004
0.05	.0030	.0026	.0022	.0019	.0017	.0016	.0010	.0007	.0006	.0005	.0004	.0003
0.10	.0026	.0022	.0019	.0016	.0015	.0013	.0009	.0006	.0005	.0004	.0003	.0002

Remember that, H_n is expected to take values close to zero when H_0 is true. Hence, H_0 will be rejected for large values of H_n. Also H_n is invariant under location-scale transformations and consistent under the assumption H_1, respectively, from Propositions 2.4 and 2.5.

4. Normality tests under evaluation

Comparison of the normality tests has received attention in the literature The goodness-of-fit tests have been discussed by many authors including Shapiro et al. (1968), Poitras (2006), Yazici and Yolacan (2007), Krauczi (2009), Romao et al. (2010), Yap and Sim (2010) and Alizadeh and Arghami (2011).

In this section we consider a large number (forty) of recent and classical statistics that have been used to test normality and in Section 5 we compare their performances with that of H_n. In the following we prefer to keep the original notation for each statistic. Concerning the notation, let $x_1, x_2, \ldots, x_n$ be a random sample of size n and $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ the corresponding order statistic. Also consider the sample mean, variance, skewness and kurtosis, defined by

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2, \qquad \sqrt{b_1} = \frac{m_3}{(m_2)^{3/2}}, \qquad b_2 = \frac{m_4}{(m_2)^2},$$

respectively, where the j-th central moment m_j is given by $m_j = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^j$ and finally consider $z_{(i)} = (x_{(i)} - \bar{x})/s$, for $i = 1, \ldots, n$.

1. Vasicek's entropy estimator (Vasicek, 1976):

$$\mathrm{KL}_{mn} = \frac{\exp\{\mathrm{HV}_{mn}\}}{s}$$

where

$$\mathrm{HV}_{mn} = \frac{1}{n}\sum_{i=1}^{n} \ln\left\{\frac{n}{2m}\left(X_{(i+m)} - X_{(i-m)}\right)\right\}, \tag{6}$$

$m < n/2$ is a positive integer and $X_{(i)} = X_{(1)}$ if $i < 1$ and $X_{(i)} = X_{(n)}$ if $i > n$. H_0 is rejected for small values of KL. Vasicek (1976) showed that the maximum power for KL was typically attained by choosing $m = 2$ for $n = 10$, $m = 3$ for $n = 20$ and $m = 4$ for $n = 50$. The lower-tail 5%-significance values of KL for $n = 10, 20$ and 50 are 2.15, 2.77 and 3.34, respectively.

2. Ebrahimi's entropy estimator (Ebrahimi, Pflughoeft and Soofi, 1994):

$$\mathrm{TE}_{mn} = \frac{\exp\{\mathrm{HE}_{mn}\}}{s},$$

where

$$\mathrm{HE}_{mn} = \frac{1}{n}\sum_{i=1}^{n} \ln\left\{\frac{n}{c_i m}\left(X_{(i+m)} - X_{(i-m)}\right)\right\}, \tag{7}$$

and $c_i = (1 + \frac{i-1}{m})\mathrm{I}_{[1,m]}(i) + 2\mathrm{I}_{[m+1,n-m]}(i) + (1 + \frac{n-i}{m})\mathrm{I}_{[n-m+1,n]}(i)$. Ebrahimi et al. (1994) proved the linear relationship between their estimator and (6). Thus for fixed values of n and m, the tests based on (6) and (7) have the same power.

3. Nonparametric distribution function of Vasicek's estimator:

$$\mathrm{TV}_{mn} = \log\sqrt{2\pi\hat{\sigma}_v^2} + 0.5 - \mathrm{HV}_{mn},$$

where HV_{mn} was defined in (6), $\hat{\sigma}_v^2 = \mathrm{Var}_{g_v}(X)$, and

$$g_v(x) = \begin{cases} 0 & x < \xi_1 \ \text{ or } \ x > \xi_{n+1}, \\ \frac{2m}{n(x_{(i+m)} - x_{(i-m)})} & \xi_i < x \leq \xi_{i+1} \ \ i = 1, \ldots, n, \end{cases}$$

where $\xi_i = \left(x_{(i-m)} + \cdots + x_{(i+m-1)}\right)/2m$. H_0 is rejected for large values of TV_{mn}. (See Park, 2003).

4. Nonparametric distribution function of Ebrahimi estimator:

$$\mathrm{TE}_{mn} = \log\sqrt{2\pi\hat{\sigma}_e^2} + 0.5 - \mathrm{HE}_{mn},$$

where HE_{mn} was defined in (7), $\hat{\sigma}_e^2 = \mathrm{Var}_{g_e}(X)$ and

$$g_e(x) = \begin{cases} 0 & x < \eta_1 \ \text{ or } \ x > \eta_{n+1} \\ \frac{1}{n(\eta_{i+1} - \eta_i)} & \eta_i < x \leq \eta_{i+1} \ \ i = 1, \ldots, n, \end{cases}$$

with

$$\eta_i = \begin{cases} \xi_{m+1} - \frac{1}{m+k-1} \sum_{k=i}^{m} (x_{(m+k)} - x_{(1)}) & 1 \le i \le m, \\ \frac{1}{2m} \left(x_{(i-m)} + \cdots + x_{(i+m-1)} \right) & m+1 \le i \le n-m+1, \\ \xi_{n-m+1} + \frac{1}{n+m-k+1} \sum_{k=n-m+2}^{i} (x_{(n)} - x_{(k-m-1)}) & n-m+2 \le i \le n+1, \end{cases}$$

and $\xi_i = \left(x_{(i-m)} + \cdots + x_{(i+m-1)} \right)/2m$. H_0 is rejected for large values of TE_{mn}. (See Park, 2003).

5. Nonparametric distribution function of Alizadeh and Arghami estimator (Alizadeh Noughabi and Arghami, 2010, 2013):

$$\mathrm{TA}_{mn} = \log \sqrt{2\pi \hat{\sigma}_a^2} + 0.5 - \mathrm{HA}_{mn},$$

where

$$\mathrm{HA}_{mn} = \frac{1}{n} \sum_{i=1}^{n} \ln \left\{ \frac{n}{a_i m} \left(X_{(i+m)} - X_{(i-m)} \right) \right\},$$

with $a_i = \mathrm{I}_{[1,m]}(i) + 2\mathrm{I}_{[m+1,n-m]}(i) + \mathrm{I}_{[n-m+1,n]}(i)$, $\hat{\sigma}_a^2 = \mathrm{Var}_{g_a}(X)$ and

$$g_a(x) = \begin{cases} 0 & x < \eta_1 \text{ or } x > \eta_{n+1}, \\ \frac{1}{n(\eta_{i+1} - \eta_i)} & \eta_i < x \le \eta_{i+1} \quad i = 1, \ldots, n, \end{cases}$$

with

$$\eta_i = \begin{cases} \xi_{m+1} - \frac{1}{m} \sum_{k=i}^{m} (x_{(m+k)} - x_{(1)}) & 1 \le i \le m, \\ \frac{1}{2m} \left(x_{(i-m)} + \cdots + x_{(i+m-1)} \right) & m+1 \le i \le n-m+1, \\ \xi_{n-m+1} + \frac{1}{m} \sum_{k=n-m+2}^{i} (x_{(n)} - x_{(k-m-1)}) & n-m+2 \le i \le n+1, \end{cases}$$

and $\xi_i = \left(x_{(i-m)} + \cdots + x_{(i+m-1)} \right)/2m$. Also $m = [\sqrt{n} + 1]$. H_0 is rejected for large values of TA_{mn}. The upper-tail 5%-significance values of TA for $n = 10, 20$ and 50 are 0.4422, 0.2805 and 0.1805, respectively.

6. Dimitriev and Tarasenko's entropy estimator (Dimitriev and Tarasenko, 1973):

$$\mathrm{TD}_{mn} = \frac{\exp\{\mathrm{HD}_{mn}\}}{s}$$

where

$$\mathrm{HD}_{mn} = -\int_{-\infty}^{\infty} \ln(\hat{f}(x))\hat{f}(x)\,dx,$$

where $\hat{f}(x)$ is the kernel density estimation of $f(x)$ given by

$$\hat{f}(X_i) = \frac{1}{nh}\sum_{j=1}^{n} k\left(\frac{X_i - X_j}{h}\right), \tag{8}$$

where k is a kernel function satisfying $\int_{-\infty}^{\infty} k(x)\,dx = 1$ and h is a bandwidth. The kernel function k being the standard normal density function and the bandwidth $h = 1.06\hat{\sigma}n^{-1/5}$. H_0 is rejected for small values of TD_{mn}.

7. Corea's entropy estimator (Corea, 1995):

$$\mathrm{TC}_{mn} = \frac{\exp\{\mathrm{HC}_{mn}\}}{s},$$

where

$$\mathrm{HC}_{mn} = -\frac{1}{n}\sum_{i=1}^{n} \ln\left\{\frac{\sum_{j=i-m}^{i+m}\left(X_{(j)} - \tilde{X}_{(i)}\right)(j-i)}{n\sum_{j=i-m}^{i+m}\left(X_{(j)} - \tilde{X}_{(i)}\right)^2}\right\}$$

and $\tilde{X}_{(i)} = \sum_{j=i-m}^{i+m} X_{(j)}/(2m+1)$. H_0 is rejected for small values of TC_{mn}.

8. Van Es's entropy estimator (Van Es, 1992):

$$\mathrm{TEs}_{mn} = \frac{\exp\{\mathrm{HEs}_{mn}\}}{s},$$

where

$$\mathrm{HEs}_{mn} = \frac{1}{n-m}\sum_{i=1}^{n-m}\left\{\ln\left(\frac{n+1}{m}(X_{(i+m)} - X_{(i)})\right)\right\} + \sum_{k=m}^{n}\frac{1}{k} + \ln(m) - \ln(n+1).$$

H_0 is rejected for small values of TEs_{mn}.

9. Zamanzade and Arghami's entropy estimator (Zamanzade and Arghami, 2012):

$$\mathrm{TZ1}_{mn} = \frac{\exp\{\mathrm{HZ1}_{mn}\}}{s},$$

where $\mathrm{HZ1}_{mn} = \frac{1}{n}\sum_{i=1}^{n} \ln(b_i)$, with

$$b_i = \frac{X_{(i+m)} - X_{(i-m)}}{\sum_{j=k_1(i)}^{k_2(i)-1} (\hat{f}(X_{(j+1)}) + \hat{f}(X_{(j)}))(X_{(j+1)} - X_{(j)})/2} \tag{9}$$

where $\hat{f}$ is defined as in (8) with the kernel function k being the standard normal density function and the bandwidth $h = 1.06\hat{\sigma}n^{-1/5}$. H_0 is rejected for small values of TZ1. For $n = 10, 20$ and 50, the lower-tail 5%-significance critical values are 3.403, 3.648 and 3.867.

10. Zamanzade and Arghami's entropy estimator (Zamanzade and Arghami, 2012):

$$\mathrm{TZ2}_{mn} = \frac{\exp\{\mathrm{HZ2}_{mn}\}}{s},$$

where $\mathrm{HZ2}_{mn} = \sum_{i=1}^{n} w_i \ln(b_i)$, being coefficients b_i's were defined in (9) and

$$w_i = \begin{cases} (m+i-1)/\sum_{i=1}^{n} w_i & 1 \le i \le m, \\ 2m/\sum_{i=1}^{n} w_i & m+1 \le i \le n-m, \qquad i = 1, \ldots, n, \\ (n-i+m)/\sum_{i=1}^{n} w_i & n-m+1 \le i \le n, \end{cases}$$

are weights proportional to the number of points used in computation of b_i's. H_0 is rejected for small values of TZ2. For $n = 10, 20$ and 50, the lower-tail 5%-significance critical values are 3.321, 3.520 and 3.721.

11. Zhang and Wu's statistics (Zhang and Wu, 2005):

$$\mathrm{Z_K} = \max_{1 \le i \le n} \left[(i-0.5)\ln\frac{i-0.5}{nF_0(Z_{(i)})} + (n-i+0.5)\ln\frac{n-i+0.5}{n(1-F_0(Z_{(i)}))} \right],$$

$$\mathrm{Z_C} = \sum_{i=1}^{n} \left(\log \frac{(1/F_0(Z_{(i)}) - 1)}{(n-0.5)/(i-0.75) - 1} \right)^2,$$

and

$$\mathrm{Z_A} = -\sum_{i=1}^{n} \left(\frac{\log F_0(Z_{(i)})}{n-i+0.5} + \frac{\log(1-F_0(Z_{(i)}))}{i-0.5} \right),$$

The null hypothesis H_0 is rejected for large values of the three test statistics.

12. Classical test statistics for normality based skewness and kurtosis from D'Agostino and Pearson (D'Agostino and Pearson, 1973):

$$\sqrt{b_1} = \frac{m_3}{(m_2)^{3/2}}, \qquad b_2 = \frac{m_4}{(m_2)^2},$$

The null hypothesis H_0 is rejected for both small and large values of the two test statistics.

13. Transformed skewness and kurtosis statistic from D'Agostino et al. (1990):

$$\mathrm{K}^2 = \left[Z(\sqrt{b_1})\right]^2 + [Z(b_2)]^2,$$

where

$$Z(\sqrt{b_1}) = \frac{\log(Y/c + \sqrt{(Y/c)^2+1})}{\sqrt{\log(w)}},$$
$$Z(b_2) = \left[\left(1 - \frac{2}{9A}\right) - \sqrt[3]{\frac{1-2/A}{1+y\sqrt{2/(A-4)}}}\right]\sqrt{\frac{9A}{2}},$$

where

$$c_1 = 6 + 8/c_2(2/c_2 + \sqrt{1+4/c_2^2}),$$

$$c_2 = (6(n^2-5n+2)/(n+7)(n+9))\sqrt{6(n+3)(n+5)/n(n-2)(n-3)},$$

$$c_3 = (b_2 - 3(n-1)/(n+1))/\sqrt{24n(n-2)(n-3)/(n+1)^2(n+3)(n+5)}.$$

and

$$Y = \sqrt{b_1}\sqrt{\frac{(n+1)(n+3)}{6(n-2)}}, \qquad w^2 = \sqrt{2\beta_2 - 1} - 1,$$
$$\beta_2 = \frac{3(n^2+27n-70)(n+1)(n+3)}{(n-2)(n+5)(n+7)(n+9)}; \quad c = \sqrt{\frac{2}{(w^2-1)}}.$$

Transformed skewness $Z(\sqrt{b_1})$ and transformed kurtosis $Z(b_2)$ is obtained by D'Agostino (1970) and Anscombe and Glynn (1983), respectively. The null hypothesis H_0 is rejected for large values of K^2.

14. Transformed skewness and kurtosis statistic by Doornik and Hansen (1994):

$$\mathrm{DH} = \left[Z(\sqrt{b_1})\right]^2 + z_2^2,$$

where

$$z_2 = \left[\left(\frac{\xi}{2a}\right)^{1/3} - 1 + \frac{1}{9a}\right]\sqrt{9a},$$

and

$$\xi = (b_2 - 1 - b_1)2k,$$

$$k = \frac{(n+5)(n+7)(n^3+37n^2+11n-313)}{12(n-3)(n+1)(n^2+15n-4)},$$

$$a = \frac{(n+5)(n+7)\left((n-2)(n^2+27n-70)+b_1(n-7)(n^2+2n-5)\right)}{6(n-3)(n+1)(n^2+15n-4)},$$

Transformed kurtosis z_2 is obtained by Shenton and Bowman (1977). The null hypothesis H_0 is rejected for large values of DH.

15. Bonett and Seier's statistic (Bonett and Seier, 2002):

$$\mathrm{Z_w} = \frac{\sqrt{n+2}(\hat{w}-3)}{3.54},$$

where $\hat{w} = 13.29\left(\ln\sqrt{m_2} - \log\left(n^{-1}\sum_{i=1}^{n}|x_i - \bar{x}|\right)\right)$. H_0 is rejected for both small and large values of $\mathrm{Z_w}$.

16. D'Agostino's statistic (D'Agostino, 1971):

$$\mathrm{D} = \frac{\sum_{i=1}^{n}(i-(n+1)/2)X_{(i)}}{n^2\sqrt{\sum_{i=1}^{n}\left(x_{(i)}-\bar{X}\right)^2}},$$

H_0 is rejected for both small and large values of D.

17. Chen and Shapiro's statistic (Chen and Shapiro, 1995):

$$\mathrm{QH} = \frac{1}{(n-1)s}\sum_{i=1}^{n-1}\frac{X_{(i+1)}-X_{(i)}}{M_{(i+1)}-M_{(i)}},$$

where $M_i = \Phi^{-1}((i-0.375)/(n+0.25))$, where Φ is the cdf of a standard normal random variable. H_0 is rejected for small values of QH.

18. Filliben's statistic (Filliben, 1975):

$$\mathrm{r} = \frac{\sum_{i=1}^{n} x_{(i)} M_{(i)}}{\sqrt{\sum_{i=1}^{n} M_{(i)}^2}\sqrt{(n-1)s^2}},$$

where $M_{(i)} = \Phi^{-1}(m_{(i)})$ and $m_{(1)} = 1-0.5^{1/n}$, $m_{(n)} = 0.5^{1/n}$ and $m_{(i)} = (i-0.3175)/(n+0.365)$ for $i = 2,\ldots,n-1$. H_0 is rejected for small values of r.

19. del Barrio et al.'s statistic (del Barrio et al., 1999):

$$\mathrm{R}_n = 1 - \frac{\left(\sum_{k=1}^{n} X_{(k)} \int_{(k-1)/n}^{k/n} F_0^{-1}(t)\, dt\right)^2}{m_2},$$

where m_2 is the sample standardized second moment. H_0 is rejected for large values of R_n.

20. Epps and Pulley statistic (Epps and Pulley, 1983):

$$\mathrm{T}_{\mathrm{EP}} = \frac{1}{\sqrt{3}} + \frac{1}{n^2}\sum_{k=1}^{n}\sum_{j=1}^{n} \exp\left(\frac{-(X_j - X_k)^2}{2m_2}\right) - \frac{\sqrt{2}}{n}\sum_{j=1}^{n} \exp\left(\frac{-(X_j - \bar{X})^2}{4m_2}\right),$$

where m_2 is the sample standardized second moment. H_0 is rejected for large values of T_{EP}.

21. Martinez and Iglewicz's statistic (Martinez and Iglewicz, 1981):

$$\mathrm{I}_n = \frac{\sum_{i=1}^{n}(X_i - M)^2}{(n-1)S_b^2},$$

where M is is the sample median and

$$S_b^2 = \frac{n\sum_{|\tilde{Z}_i|<1}(X_i - M)^2(1-\tilde{Z}_i^2)^4}{\left(\sum_{|\tilde{Z}_i|<1}(1-\tilde{Z}_i^2)(1-5\tilde{Z}_i^2)\right)^2},$$

with $\tilde{Z}_i = (X_i - M)/(9A)$ for $|\tilde{Z}_i| < 1$ and $\tilde{Z}_i = 0$ otherwise, and A is the median of $|X_i - M|$. H_0 is rejected for large values of I_n.

22. deWet and Venter statistic (de Wet and Venter, 1972):

$$\mathrm{E}_n = \sum_{i=1}^{n} \left(X_{(i)} - \bar{X} - s\Phi^{-1}\left(\frac{i}{n+1}\right)\right)^2 \Big/ s^2.$$

H_0 is rejected for large values of $\mathrm{E_n}$.

23. Optimal test (Csörgo and Révész, 1971):

$$\mathrm{M}_n = \sum_{i=1}^{n} \left(X_{(i)} - \bar{X} - s\Phi^{-1}\left(\frac{i}{n+1}\right)\right)^2 \phi\left(\Phi^{-1}\left(\frac{i}{n+1}\right)\right)\left[\Phi^{-1}\left(\frac{i}{n+1}\right)\right]^{\lambda-1}.$$

H_0 is rejected for large values of M_n.

24. Pettitt statistic (Pettitt, 1977):

$$\mathrm{Q}_n = \sum_{i=1}^{n} \left(\Phi\left(\frac{X_{(i)} - \bar{X}}{s}\right) - \frac{i}{n+1}\right)^2 \left[\phi\left(\Phi^{-1}\left(\frac{i}{n+1}\right)\right)\right]^{-2}.$$

H_0 is rejected for large values of $\mathrm{Q_n}$.

25. Three test statistics from LaRiccia (1986):

$$\mathrm{T}_{1n} = C_{1n}^2/(s^2 B_{1n}), \qquad \mathrm{T_{2n}} = C_{2n}^2/(s^2 B_{2n}), \qquad \mathrm{T_{3n}} = \mathrm{T_{1n}} + \mathrm{T_{2n}},$$

where

$$C_{1n} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[W_1\left(\frac{i}{n+1}\right) - A_{1n}\right] X_{(i)},$$
$$C_{2n} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[W_2\left(\frac{i}{n+1}\right) - A_{2n}\Phi^{-1}\left(\frac{i}{n+1}\right)\right] X_{(i)},$$

Also $W_1(u) = [\Phi^{-1}(u)]^2 - 1$ and $W_2(u) = [\Phi^{-1}(u)]^3 - 3\Phi^{-1}(u)$. The constants A_{1n}, A_{2n}, B_{1n} and B_{2n} are given in Table 1 from LaRiccia (1986). For all three statistics H_0 is rejected for large value.

26. Kolmogorov-Smirnov's (Lilliefors) statistic (Kolmogorov, 1933):

$$\mathrm{KS} = \max\left\{ \max_{1\le j\le n}\left[\frac{j}{n} - F_0(Z_{(j)})\right], \max_{1\le j\le n}\left[F_0(Z_{(j)}) - \frac{j-1}{n}\right]\right\}.$$

Lilliefors (1967) computed estimated critical points for the Kolmogorov-Smirnov's test statistic for testing normality when mean and variance estimated.

27. Kuiper's statistic (Kuiper, 1962):

$$\mathrm{V} = \max_{1\le j\le n}\left[\frac{j}{n} - F_0(Z_{(j)})\right] + \max_{1\le j\le n}\left[F_0(Z_{(j)}) - \frac{j-1}{n}\right].$$

Louter and Kort (1970) computed estimated critical points for the Kuiper test statistic for testing normality when mean and variance estimated.

28. Cramér-von Mises' statistic (Cramér, 1928 and von Mises, 1931):

$$\mathrm{W}^2 = \frac{1}{12n} + \sum_{j=1}^{n}\left(F_0(Z_{(j)}) - \frac{2j-1}{2n}\right)^2.$$

29. Watson's statistic (Watson, 1961):

$$\mathrm{U}^2 = \mathrm{W}^2 - n\left(\frac{1}{n}\sum_{j=1}^{n} F_0(Z_{(j)}) - \frac{1}{2}\right)^2.$$

30. Anderson-Darling's statistic (Anderson, 1954):

$$\mathrm{A}^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left(\log(F_0(Z_{(i)})) + \log\left(1 - F_0(Z_{(n-i+1)})\right)\right).$$

These classical tests are based on the empirical distribution function and H_0 is rejected for large values of KS, V, W^2, U^2 and A^2.

31. Pearson's chi-square statistic (D'Agostino and Stephens, 1986):

$$\mathrm{P} = \sum_i (C_i - E_i)^2/E_i,$$

where C_i is the number of counted and E_i is the number of expected observations (under H_0) in class i. The classes are build is such a way that they are equiprobable under the null hypothesis of normality. The number of classes used for the test is $\lceil 2n^{2/5}\rceil$ where $\lceil . \rceil$ is ceiling function.

32. Shapiro-Wilk's statistic (Shapiro and Wilk, 1965):

$$\mathrm{SW}=\frac{\left(\sum_{i=1}^{[n/2]} a_{(n-i+1)}\left(X_{(n-i+1)}-X_{(i)}\right)\right)^2}{\sum_{i=1}^{n}\left(X_{(i)}-\bar{X}\right)^2},$$

where coefficients a_i's are given by

$$(a_1,\ldots,a_n)=\frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}}, \tag{10}$$

and $m^T=(m_1,\ldots,m_n)$ and V are, respectively, the vector of expected values and the covariance matrix of the order statistic of n iid random variables sampled from the standard normal distribution. H_0 is rejected for small values of SW.

33. Shapiro-Francia's statistic (Shapiro and Francia, 1972) is a modification of SW. It is defined as

$$\mathrm{SF}=\frac{\left(\sum_{i=1}^{n} b_i X_{(i)}\right)^2}{\sum_{i=1}^{n}(X_{(i)}-\bar{X})^2},$$

where

$$(b_1,\ldots,b_n)=\frac{m^T}{(m^T m)^{1/2}}$$

and m is defined as in (10). H_0 is rejected for small values of SF.

34. SJ statistic discussed in Gel, Miao and Gastwirth (2007). It is based on the ratio of the classical standard deviation $\hat{\sigma}$ and the robust standard deviation J_n (average absolute deviation from the median (MAAD)) of the sample data

$$\mathrm{SJ}=\frac{s}{J_n}, \tag{11}$$

where $J_n=\sqrt{\frac{\pi}{2}}\frac{1}{n}\sum_{i=1}^{n}|X_i-M|$ and M is the sample median. H_0 is rejected for large values of SJ.

35. Jarque-Bera's statistic (Jarque and Bera, 1980, 1987):

$$\mathrm{JB}=\frac{n}{6}b_1+\frac{n}{24}\left(b_2-3\right)^2,$$

where $\sqrt{b_1}$ and b_2 are the sample skewness and sample kurtosis, respectively. H_0 is rejected for large values of JB.

36. Robust Jarque-Bera's statistic (Gel and Gastwirth, 2008):

$$\mathrm{RJB} = \frac{n}{C_1}\left(\frac{m_3}{J_n^3}\right)^2 + \frac{n}{C_2}\left(\frac{m_4}{J_n^4} - 3\right)^2,$$

where J_n is defined as in (11), C_1 and C_2 are positive constants. For a 5%-significance level, $C_1 = 6$ and $C_2 = 64$ according to Monte Carlo simulations. H_0 is rejected for large values of RJB.

5. Simulation study

In this section we study the power of the normality test based on $\mathrm{H_n}$ and compare it with a large number of recent and classical normality tests. To facilitate comparisons of the power of the present test with the powers of the mentioned tests, we select two sets of alternative distributions:

Set 1. Alternatives listed in in Esteban et al. (2001).

Set 2. Alternatives listed in Gan and Koehler (1990) and Krauczi (2009).

Set 1 of alternative distributions

Following Esteban et al. (2001) we consider the following alternative distributions, that can be classified in four groups:

Group I: Symmetric distributions with support on $(-\infty, \infty)$:

- Standard Normal (N);
- Student's t (t) with 1 and 3 degrees of freedoms;
- Double Exponential (DE) with parameters $\mu = 0$ (location) and $\sigma = 1$ (scale);
- Logistic (L) with parameters $\mu = 0$ (location) and $\sigma = 1$ (scale);

Group II: Asymmetric distributions with support on $(-\infty, \infty)$:

- Gumbel (Gu) with parameters $\alpha = 0$ (location) and $\beta = 1$ (scale);
- Skew Normal (SN) with with parameters $\mu = 0$ (location), $\sigma = 1$ (scale) and $\alpha = 2$ (shape);

Group III: Distributions with support on $(0,\infty)$:

- Exponential (Exp) with mean 1;
- Gamma (G) with parameters $\beta = 1$ (scale) and $\alpha = .5, 2$ (shape);
- Lognormal (LN) with parameters $\mu = 0$ and $\sigma = .5, 1, 2$;
- Weibull (W) with parameters $\beta = 1$ (scale) and $\alpha = .5, 2$ (shape);

Group IV: Distributions with support on $(0,1)$:

- Uniform (Unif);
- Beta (B) with parameters (2,2), (.5,.5), (3,1.5) and (2,1).

Set 2 of alternative distributions

Gan and Koehler (1990) and Krauczi (2009) considered a battery of "difficult alternatives" for comparing normality tests. We also consider them in order to evaluate the sensitivity of the proposed test. Let U and Z denote a $[0,1]$-Uniform and a Standard Normal random variable, respectively.

- Contaminated Normal distribution (CN) with parameters $(\lambda, \mu_1, \mu_2, \sigma)$ given by the cdf $F(x) = (1-\lambda)F_0(x, \mu_1, 1) + \lambda F_0(x, \mu_2, \sigma)$;
- Half Normal (HN) distribution, that is, the distribution of $|Z|$.
- Bounded Johnson's distribution (SB) with parameters (γ, δ) of the random variable $e^{(Z-\gamma)/\delta}/(1+e^{(Z-\gamma)/\delta})$;
- Unbounded Johnson's distribution (UB) with parameters (γ, δ) of the random variable $\sinh((Z-\gamma)/\delta)$;
- Triangle type I (Tri) with density function $f(x) = 1-|t|$, $-1 < t < 1$;
- Truncated Standard Normal distribution at a and b (TN);
- Tukey's distribution (Tu) with parameter λ of the random variable $U^\lambda - (1-U)^\lambda$.
- Cauchy distribution with parameters $\mu = 0$ (location), $\sigma = 1$ (scale).
- Chi-squared distribution χ^2 with k degrees of freedom.

Tables 2-3 contain the skewness $(\sqrt{\beta_1})$ and kurtosis (β_2) of the previous sets of alternative distributions. Alternatives in *Set 2* are roughly ordered and grouped in five groups according to their skewness and kurtosis values in Table 3. These groups correspond to: symmetric short tailed, symmetric closed to normal, asymmetric short tailed, asymmetric long tailed. Figure 2 illustrates some of the possible shapes of the pdf's of the alternatives in *Set 1* and *Set 2*.

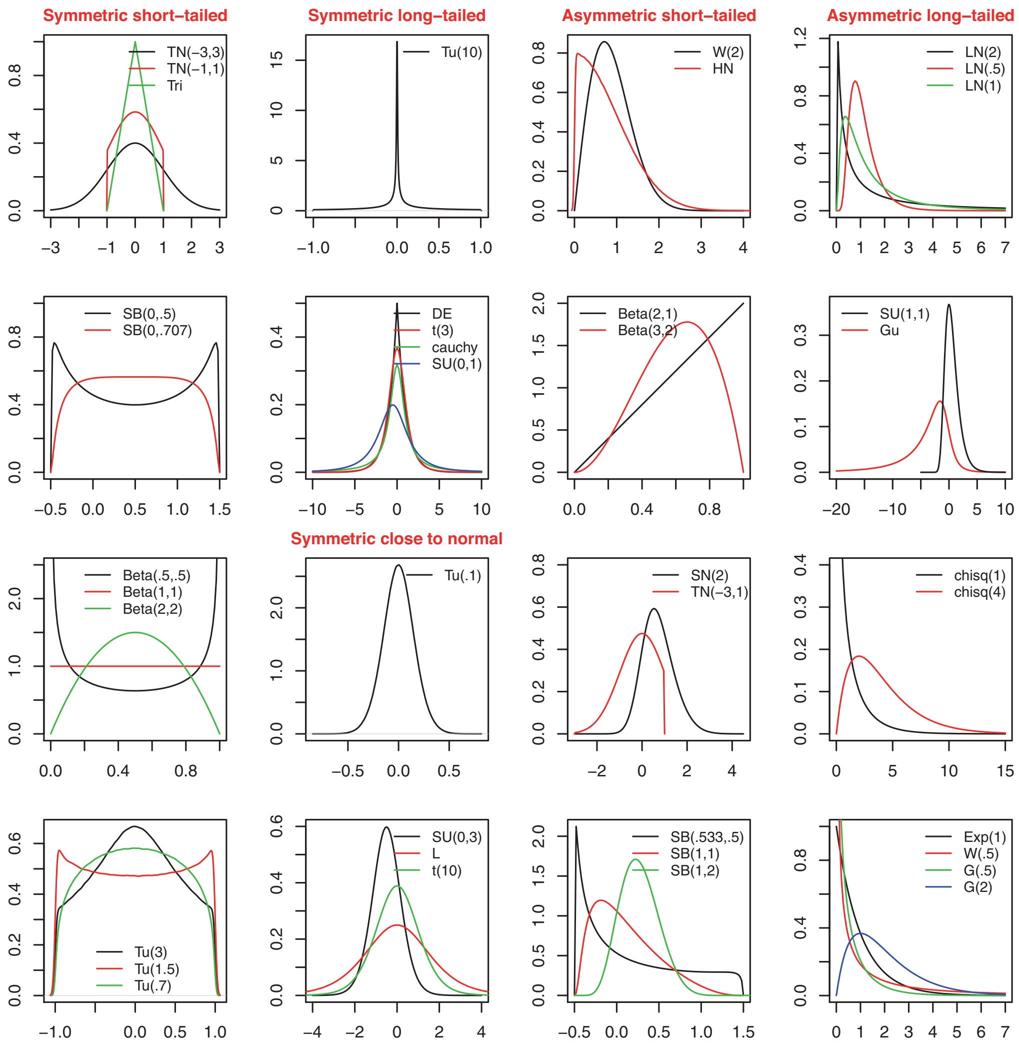

Figure 2: *Plots of alternative distributions in* Set 1 *and* Set 2.

Tables 4-5 contain the estimated value of $\mathrm{H_n}$ (for $h(x) = (x-1)^2/(x+1)^2$ and $h(x) = x\log(x) - x + 1$, respectively), for each alternative distribution, computed as the average value from 10000 simulated samples of sizes $n = 10, 20, 50, 100, 1000$. In the last row of these tables $(n = \infty)$), we show the value of $D(F_0, F_1)$ computed with the the command `integrate` in R Software, with (μ) and (σ^2) being the expectation and variance of F_1, respectively. *These tables show consistency of the test statistic* H_n.

Tables 6-7 report the power of the 5% significance level of forty normality tests based on the statistics considered in Section 4 for the *Set 1* of alternatives.

Tables 8-9 contain the power of the 5% significance level test of normality based on the most powerful statistics and the alternatives listed in *Set 2*.

Table 2: *Skewness and kurtosis of alternative distributions in* Set 1.

	Group I				Group II		Group III								Group IV				
	t(1)	t(3)	L	DE	Gu	SN(2)	Exp	G(2)	G(.5)	LN(1)	LN(2)	LN(.5)	W(.5)	W(2)	Unif	B(2,2)	B(.5,.5)	B(3,.5)	B(2,1)
$\sqrt{\beta_1}$	0	0	0	0	1.30	.45	2	1.41	2.83	6.18	414.36	1.75	6.62	.63	0	0	0	−1.575	−.57
β_2	—	—	4.2	6	5.4	.31	9	6	15	113.94	9220560	8.90	87.72	3.25	1.8	2.14	1.5	5.22	2.4

Table 3: *Skewness and kurtosis of alternative distributions in* Set 2.

	Symmetric													Asymmetric							
	Short tailed							Close to Normal			Long tailed			Short tailed					Long tailed		
	Tu	Tu	Tu	SB	Tri	TN	TN	Tu	SU	t	Tu	SU	caushy	TN	SB	SB	SB	HN	SU	χ^2	χ^2
	(.7)	(1.5)	(3)	(0,.5)		(−1,1)	(−3,3)	(.1)	(0,3)	(10)	(10)	(0,1)		(−3,1)	(1,1)	(1,2)	(.533,.5)		(1,1)	(1)	(4)
$\sqrt{\beta_1}$	0	0	0	0	0	0	0	0	0	0	0	0	0	−.55	.73	.28	.65	.97	−5.37	2.83	1.41
β_2	1.92	1.75	2.06	1.63	2.4	1.94	2.84	3.21	3.53	4	5.38	36.2	∞	2.78	2.91	2.77	2.13	3.78	93.4	15	6

Table 4: *Estimated value of* H_n *with* $h_1(x) = (x-1)^2/(x+1)^2$ *under* H_1*, based on 10000 simulations for several values of n.*

	Group I				Group II		Group III								Group IV				
n	t(1)	t(3)	L	DE	Gu	SN(2)	Exp	G(2)	G(.5)	LN(1)	LN(2)	LN(.5)	W(.5)	W(2)	Unif	B(2,2)	B(.5,.5)	B(3,.5)	B(2,1)
10	—	.0011	.00086	.0010	.0011	.00092	.0017	.0013	.0025	.00226	.0040	.0013	.0035	.00097	.0009	.00082	.0012	.0013	.0008
20	—	.0007	.00043	.0006	.0007	.00047	.0014	.0009	.0023	.00213	.0045	.0009	.0036	.00054	.0005	.00041	.0008	.0011	.0005
50	—	.0005	.00018	.0004	.0004	.00022	.0012	.0006	.0022	.00211	.0052	.0007	.0037	.00028	.0003	.00019	.0006	.0011	.0003
100	—	.0004	.00011	.0003	.0003	.00013	.0011	.0006	.0022	.00215	.0056	.0006	.0039	.00019	.0003	.00012	.0006	.0011	.0003
1000	—	.0004	.00004	.0002	.0002	.00006	.0010	.0005	.0021	.00226	.0066	.0005	.0040	.00012	.0002	.00006	.0005	.0011	.0002
∞	—	.0004	.00003	.0002	.0002	.00005	.0010	.0005	.0021	.00228	.0074	.0005	.0040	.00011	.0002	.00006	.0005	.0011	.0002

Table 5: *Estimated value of* H_n *with* $h_2(x) = x\log(x) - x + 1$ *under* H_1*, based on 10000 simulations for several values of n.*

	Group I				Group II		Group III								Group IV				
n	t(1)	t(3)	L	DE	Gu	SN(2)	Exp	G(2)	G(.5)	LN(1)	LN(2)	LN(.5)	W(.5)	W(2)	Unif	B(2,2)	B(.5,.5)	B(3,.5)	B(2,1)
10	—	.0021	.00167	.0019	.0022	.0018	.0034	.0027	.0048	.0044	.0077	.0026	.0065	.0020	.0019	.0017	.0025	.0027	.0017
20	—	.0014	.00086	.0012	.0013	.0009	.0028	.0017	.0045	.0042	.0088	.0018	.0070	.0010	.0011	.0009	.0017	.0024	.0010
50	-	.0010	.00037	.0007	.0008	.0004	.0023	.0013	.0044	.0042	.0106	.0013	.0075	.0005	.0006	.0004	.0013	.0023	.0006
100	-	.0009	.00021	.0006	.0006	.0003	.0022	.001	.0043	.0043	.0113	.0012	.0079	.0004	.0005	.0003	.0012	.0023	.0005
1000	—	.0009	.00007	.0004	.0005	.0001	.0021	.0009	.0043	.0046	.0139	.0010	.0084	.0002	.0004	.0001	.0011	.0023	.0004
∞	—	.0009	.00006	.0004	.0005	.0001	.0021	.0009	.0043	.0047	.0163	.0010	.0084	.0002	.0004	.0001	.0010	.0023	.0004

Table 6: *Power comparisons for the normality test for* Set 1 *of alternative distributions,* $\alpha = 0.05$, $n = 10$.

	Group		I				II		III								IV				
	altern.	N	t(1)	t(3)	L	DE	Gu	SN	Exp	G(2)	G(.5)	LN(1)	LN(2)	LN(.5)	W(.5)	W(2)	Unif	B(2,2)	B(.5,.5)	B(3,.5)	B(2,1)
1	KL	.048	.442	.091	.051	.091	.101	.058	.416	.179	.782	.552	.938	.181	.931	.075	.167	.082	.512	.108	.173
2	TV	.048	.375	.082	.048	.053	.092	.055	.397	.151	.762	.519	.933	.144	.923	.073	**.181**	.084	**.514**	.656	.170
3	TE	.052	.460	.112	.058	.077	.111	.059	.454	.185	.794	.581	.945	.181	.935	.074	.158	.071	.481	.686	.164
4	TA	.053	.507	.134	.065	.094	.124	.062	.477	.213	**.810**	.616	**.951**	.208	**.940**	.080	.129	.064	.451	**.704**	.162
5	TD	.051	.583	.201	.087	.163	.154	.071	.394	.222	.631	.565	.869	.249	.813	.076	.028	.025	.080	.065	.093
6	TC	.054	.409	.083	.047	.057	.097	.053	.404	.173	.786	.542	.936	.171	.926	.071	.170	**.086**	.489	.110	**.182**
7	TEs	.049	.591	.167	.074	.140	.113	.062	.330	.158	.679	.485	.892	.176	.876	.064	.061	.037	.238	.064	.092
8	TZ1	.053	.632	.212	.089	.177	.145	.068	.359	.209	.581	.524	.846	.229	.784	.074	.030	.025	.078	.061	.081
9	TZ2	.051	.638	.216	.091	.181	.144	.066	.353	.205	.572	.516	.840	.228	.776	.073	.026	.023	.060	.058	.076
10	Z_K	.055	.587	.174	.075	.154	.126	.071	.352	.180	.636	.509	.885	.192	.842	.079	.078	.053	.221	.510	.109
11	Z_C	.053	.580	.183	.079	.154	.157	.074	.450	.245	.740	.606	.926	.248	.898	.089	.094	.044	.336	.621	.130
12	Z_A	.053	.608	.199	.083	.167	.162	.071	.457	.246	.744	.612	.928	.255	.901	.086	.050	.032	.204	.621	.115
13	$\sqrt{b_1}$	.057	.587	.219	.096	.184	.165	.073	.372	.226	.557	.532	.928	.247	.751	.088	.019	.024	.035	.437	.083
14	b_2	.053	.536	.170	.073	.136	.113	.060	.227	.148	.340	.353	.907	.159	.508	.072	.115	.057	.270	.235	.092
15	K^2	.058	.592	.220	.096	.190	.154	.073	.314	.197	.467	.464	.754	.221	.662	.082	.020	.021	.065	.336	.067
16	DH	.055	.625	.207	.084	.183	.130	.067	.344	.183	.590	.507	.860	.195	.797	.069	.071	.037	.238	.467	.093
17	Z_w	.055	.501	.150	.068	.130	.075	.088	.125	.091	.181	.210	.416	.097	.311	.055	.100	.056	.215	.123	.073
18	D	.051	.584	.175	.071	.142	.111	.060	.270	.146	.478	.434	.799	.168	.717	.064	.042	.044	.039	.335	.061
19	QH	.053	.598	.189	.081	.159	.160	.075	.455	.245	.742	.609	.928	.250	.901	.090	.094	.046	.321	.625	.135
20	r	.054	.635	.214	.088	.187	.160	.074	.421	.231	.692	.578	.907	.245	.868	.089	.042	.031	.164	.561	.099
21	R_n	.054	.609	.196	.083	.167	.162	.075	.448	.244	.733	.604	.924	.251	.894	.090	.077	.042	.276	.613	.125
22	T_{EP}	.053	.602	.200	.088	.170	.167	.077	.427	.244	.663	.587	.891	.256	.842	.070	.054	.040	.152	.538	.115
23	I_n	.055	.157	.151	.084	.151	.120	.066	.209	.149	.199	.215	.100	.151	.134	.070	.024	.025	.043	.207	.065
24	E_n	.055	.638	.218	.089	.193	.158	.073	.407	.226	.670	.567	.898	.240	.852	.082	.035	.028	.126	.536	.091
25	M_n	.054	.631	.226	.095	.198	.147	.071	.326	.189	.524	.484	.808	.214	.733	.073	.014	.020	.029	.385	.061
26	Q_n	.053	.604	.175	.074	.152	.141	.071	.426	.220	.728	.585	.923	.222	.894	.081	.094	.051	.285	.610	.130
27	T_{1n}	.054	.516	.179	.083	.145	.173	.072	.475	.264	.726	.626	.918	.274	.884	.095	.036	.030	.093	.605	.114
28	T_{2n}	.053	.555	.168	.072	.155	.075	.055	.106	.075	.167	.206	.453	.088	.326	.049	.090	.046	.284	.105	.060
29	T_{3n}	.057	.647	.225	.093	.204	.146	.070	.360	.199	.625	.518	.882	.216	.831	.074	.039	.026	.203	.487	.076
30	KS	.053	.581	.164	.073	.148	.124	.072	.312	.170	.545	.469	.828	.182	.761	.078	.066	.051	.163	.424	.103
31	V	.050	.593	.163	.071	.143	.119	.065	.365	.180	.662	.530	.894	.188	.856	.074	.087	.054	.240	.540	.108
32	W^2	.052	.624	.186	.080	.164	.143	.073	.396	.210	.674	.562	.898	.220	.860	.082	.083	.050	.236	.552	.116
33	U^2	.052	.618	.178	.076	.159	.135	.071	.382	.200	.661	.547	.893	.211	.853	.081	.091	.056	.260	.540	.120
34	A^2	.051	.619	.190	.083	.165	.147	.073	.417	.225	.670	.578	.911	.233	.877	.085	.086	.048	.268	.580	.126
35	P	.042	.531	.148	.083	.136	.127	.080	.397	.200	.704	.545	.903	.199	.878	.087	.086	.061	.229	.594	.136
36	SW	.052	.597	.187	.082	.159	.159	.075	.451	.245	.740	.608	.927	.248	.899	.088	.090	.045	.312	.622	.133
37	SF	.054	.631	.214	.088	.185	.161	.074	.426	.234	.701	.584	.912	.248	.872	.085	.047	.033	.183	.571	.104
38	SJ	.055	**.655**	.217	.096	**.211**	.121	.068	.253	.147	.429	.416	.756	.176	.660	.060	.012	.021	.022	.285	.046
39	JB	.059	.600	.223	.096	.192	.149	.075	.352	.219	.532	.511	.804	.242	.731	.087	.016	.021	.029	.396	.073
40	RJB	.056	.644	**.228**	**.097**	.205	.165	.072	.485	.189	.504	.470	.784	.214	.700	.076	.015	.021	.025	.345	.061
h_2	H_n	.051	.596	.173	.074	.150	.190	.091	.504	.285	.780	.659	.940	.290	.918	.114	.074	.046	.218	.331	.054
h_1	H_n	.051	.587	.169	.073	.144	**.199**	**.095**	**.516**	**.296**	.784	**.665**	.942	**.301**	.920	**.119**	.079	.049	.220	.300	.047

Table 7: *Power comparisons for the normality test for* Set 1 *of alternative distributions,* $\alpha = 0.05$, $n = 20$.

	Group		I				II		III								IV				
	altern.	N	t(1)	t(3)	L	DE	Gu	SN	Exp	G(2)	G(.5)	LN(1)	LN(2)	LN(.5)	W(.5)	W(2)	Unif	B(2,2)	B(.5,.5)	B(3,.5)	B(2,1)
1	KL	.045	.737	.165	.051	.091	.198	.073	.846	.457	.992	.927	.999	.404	**1.00**	.132	.442	.131	.914	.224	**.438**
2	TV	.047	.684	.121	.046	.062	.176	.067	.830	.429	.992	.910	**1.00**	.364	**1.00**	.126	**.443**	**.136**	**.910**	.980	.428
3	TE	.047	.786	.205	.064	.129	.237	.079	.865	.508	**.993**	.934	**1.00**	.445	**1.00**	.143	.391	.112	.891	**.984**	.423
4	TA	.048	.858	.301	.095	.229	.279	.101	**.870**	.533	**.993**	**.937**	**1.00**	.485	**1.00**	.145	.258	.064	.824	.983	.358
5	TD	.049	.872	.371	.134	.304	.310	.102	.790	.507	.959	.909	.997	.517	.995	.148	.084	.028	.408	.129	.221
6	TC	.047	.687	.138	.043	.070	.185	.076	.836	.443	.991	.919	.999	.386	.999	.133	.438	.135	.902	.225	.432
7	TEs	.054	.871	.330	.114	.271	.195	.073	.646	.322	.955	.825	.997	.360	.997	.089	.076	.027	.460	.069	.131
8	TZ1	.056	.885	.377	.133	.309	.294	.099	.745	.459	.947	.895	.996	.470	.994	.123	.099	.028	.442	.114	.200
9	TZ2	.062	.900	.402	.147	.344	.282	.096	.688	.416	.915	.865	.994	.445	.987	.110	.028	.013	.145	.079	.130
10	Z_K	.055	.861	.308	.109	.252	.251	.088	.797	.438	.983	.906	.992	.423	.999	.118	.132	.054	.512	.952	.253
11	Z_C	.050	.844	.333	.121	.249	.313	.104	.838	.529	.983	.931	.999	.520	.999	.159	.231	.052	.782	.953	.307
12	Z_A	.052	.864	.347	.124	.268	.323	.108	.866	**.559**	.989	.943	.999	.541	.999	.166	.142	.032	.674	.967	.318
13	$\sqrt{b_1}$	.052	.775	.345	.135	.286	.324	.114	.708	.471	.891	.869	.990	.508	.979	.151	.006	.008	.013	.762	.125
14	b_2	.049	.832	.333	.111	.239	.181	.076	.365	.230	.544	.600	.877	.279	.787	.093	.324	.109	.683	.316	.122
15	K^2	.048	.849	.370	.139	.282	.267	.100	.570	.371	.777	.781	.967	.418	.936	.119	.133	.030	.491	.587	.093
16	DH	.050	.871	.382	.141	.316	.258	.089	.730	.429	.941	.888	.997	.444	.994	.110	.101	.024	.494	.855	.186
17	Z_w	.049	.853	.326	.108	.280	.120	.062	.203	.135	.340	.427	.756	.173	.602	.059	.225	.089	.539	.160	.111
18	D	.051	.882	.347	.119	.276	.202	.075	.517	.280	.805	.758	.984	.330	.963	.086	.094	.075	.031	.607	.067
19	QH	.053	.862	.327	.115	.251	.313	.103	.841	.533	.983	.933	.999	.520	.999	.157	.229	.059	.761	.957	.326
20	r	.053	.895	.389	.145	.325	.311	.108	.794	.492	.970	.911	.998	.504	.999	.145	.073	.019	.460	.916	.207
21	R_n	.054	.875	.353	.128	.281	.320	.108	.833	.528	.981	.931	.999	.524	.999	.158	.176	.045	.683	.946	.292
22	T_{EP}	.054	.868	.332	.115	.257	.309	.104	.778	.502	.954	.912	.998	.507	.995	.147	.130	.043	.478	.888	.266
23	I_n	.053	.144	.268	.145	.286	.216	.091	.387	.289	.286	.310	.038	.313	.084	.095	.004	.006	.013	.382	.070
24	E_n	.053	.901	.398	.150	.337	.302	.105	.763	.467	.959	.899	.998	.488	.997	.135	.038	.013	.326	..891	.169
25	M_n	.050	.894	.409	.153	.339	.274	.099	.661	.395	.897	.841	.992	.431	.984	.112	.005	.004	.025	.771	.087
26	Q_n	.053	.874	.311	.105	.257	.277	.092	.847	.508	.988	.932	.999	.486	.999	.142	.176	.051	.663	.963	.333
27	T_{1n}	.050	.656	.255	.106	.179	**.345**	.111	.838	.569	.972	.938	.999	**.565**	.999	.178	.029	.018	.082	.924	.246
27	T_{2n}	.049	.866	.343	.116	.296	.100	.059	.150	.101	.269	.362	.734	.141	.554	.049	.311	.079	.773	.109	.119
27	T_{3n}	.050	.897	.387	.143	.330	.278	.096	.779	.453	.973	.905	.999	.466	.999	.121	.174	.032	.732	.926	.225
30	KS	.056	.847	.268	.089	.227	.214	.084	.595	.338	.884	.799	.992	.349	.985	.109	.102	.056	.377	.761	.192
31	V	.052	.863	.273	.090	.236	.199	.073	.697	.352	.955	.859	.998	.348	.997	.093	.148	.063	.495	.885	.205
32	W^2	.056	.880	.308	.105	.265	.254	.091	.732	.420	.954	.883	.998	.429	.996	.123	.149	.056	.517	.882	.237
33	U^2	.055	.878	.297	.099	.261	.225	.083	.694	.381	.942	.862	.997	.391	.995	.113	.167	.064	.554	.863	.230
34	A^2	.054	.880	.324	.110	.268	.279	.094	.780	.463	.968	.906	.999	.467	.998	.133	.179	.056	.624	.917	.269
35	P	.049	.777	.182	.067	.144	.141	.063	.656	.282	.956	.827	.998	.267	.994	.074	.082	.053	.272	.880	.162
36	SW	.054	.867	.337	.119	.266	.317	.105	.840	.534	.982	.933	.999	.526	.999	.160	.208	.053	.738	.954	.314
37	SF	.053	.893	.383	.143	.318	.313	.107	.802	.498	.973	.915	.998	.507	.999	.148	.086	.022	.499	.922	.220
38	SJ	.053	**.915**	.404	.147	**.377**	.188	.078	.416	.234	.676	.695	.959	.301	.911	.064	.003	.006	.002	.435	.033
39	JB	.050	.864	.384	.146	.300	.285	.104	.630	.404	.840	.825	.984	.448	.964	.125	.003	.004	.006	.677	.080
40	RJB	.054	.906	**.410**	**.159**	.354	.266	.099	.563	.357	.790	.787	.977	.410	.949	.106	.002	.004	.004	.594	.061
h_2	H_n	.055	.874	.302	.100	.254	.322	.116	.832	.540	.982	.933	.999	.525	.999	.176	.154	.061	.524	.734	.140
h_1	H_n	.053	.869	.293	.097	.244	.330	**.120**	.835	.546	.983	.934	.999	.532	.999	**.181**	.156	.062	.525	.709	.126

Table 8: *Power comparisons for the normality test for* Set 2 *of alternative distributions,* $\alpha = 0.05$, $n = 10$.

	Symmetric													Asymmetric							
	Short tailed							Close to Normal			Long tailed			Short tailed					Long tailed		
	Tu	Tu	Tu	SB	Tri	TN	TN	Tu	SU	t	Tu	SU	caushy	TN	SB	SB	SB	HN	SU	χ^2	χ^2
	(.7)	(1.5)	(3)	(0,.5)		(-1,1)	(-3,3)	(.1)	(0,3)	(10)	(10)	(0,1)		(-3,1)	(1,1)	(1,2)	(.533,.5)		(1,1)	(1)	(4)
TV	**.122**	**.205**	**.095**	**.313**	**.057**	**.125**	**.053**	.044	.045	.057	.281	.086	.369	.027	.130	.055	**.442**	.206	.260	.766	.171
TA	.086	.155	.068	.244	.043	.089	.047	.050	.051	.048	.460	.162	.496	**.095**	.143	.051	.431	.177	.357	**.814**	.225
Z_A	.040	.060	.033	.101	.031	.037	.043	.062	.067	.080	.523	.258	.608	.071	.133	.050	.296	.196	.438	.755	.254
$\sqrt{b_1}$	.018	.019	.017	.024	.025	.017	.040	.061	.069	.084	.372	.255	.574	.061	.101	.042	.130	.156	.412	.566	.221
r	.035	.050	.021	.086	.031	.034	.044	.064	.068	.087	.595	.276	.634	.070	.123	.050	.253	.177	.436	.704	.239
R_n	.057	.094	.042	.149	.035	.054	.046	.061	.066	.075	.552	.251	.609	.073	.142	.054	.323	.194	.436	.743	.251
T_{EP}	.046	.067	.037	.101	.035	.047	.050	.063	.068	.085	.480	.180	.602	.072	.143	.057	.271	.193	**.448**	.684	.256
E_n	.029	.038	.026	.067	.030	.028	.044	.063	.068	.090	.602	.278	.639	.067	.114	.048	.224	.168	.433	.682	.232
M_n	.015	.017	.017	.020	.043	.016	.045	.064	**.075**	.100	.521	.287	.634	.061	.087	.044	.115	.137	.397	.550	.200
T_{1n}	.032	.041	.026	.060	.031	.030	.044	.061	.063	.069	.338	.220	.517	.077	.140	.051	.253	.204	.441	.739	.266
T_{3n}	.029	.047	.026	.088	.030	.030	.045	.064	.072	.038	.579	.288	.645	.083	.093	.046	.200	.139	.405	.644	.198
A^2	.063	.102	.049	.154	.037	.063	.050	.060	.064	.068	.630	.245	.620	.075	.137	.054	.319	.182	.422	.715	.234
SW	.064	.109	.049	.170	.036	.064	.046	.060	.064	.074	.532	.242	.598	.078	.144	.054	.345	.199	.433	.751	.253
SF	.037	.055	.031	.093	.030	.036	.044	.063	.067	.082	.588	.270	.630	.070	.124	.050	.261	.179	.435	.709	.238
SJ	.014	.016	.019	.015	.032	.016	.047	**.067**	.072	**.093**	**.678**	.290	**.660**	.049	.070	.046	.086	.101	.360	.442	.157
RJB	.015	.016	.016	.018	.026	.016	.045	.064	.073	**.093**	.569	**.290**	.645	.057	.088	.044	.105	.132	.394	.504	.195
H_n	.066	.094	.052	.141	.047	.061	**.053**	.060	.064	.060	.625	.222	.592	.026	**.208**	**.073**	.416	**.262**	.264	.807	**.315**

Table 9: *Power comparisons for the normality test for* Set 2 *of alternative distributions,* $\alpha = 0.05$, $n = 20$.

	Symmetric													Asymmetric							
	Short tailed							Close to Normal			Long tailed			Short tailed					Long tailed		
	Tu	Tu	Tu	SB	Tri	TN	TN	Tu	SU	t	Tu	SU	caushy	TN	SB	SB	SB	HN	SU	χ^2	χ^2
	(.7)	(1.5)	(3)	(0,.5)		(-1,1)	(-3,3)	(.1)	(0,3)	(10)	(10)	(0,1)		(-3,1)	(1,1)	(1,2)	(.533,.5)		(1,1)	(1)	(4)
TV	**.291**	**.515**	**.188**	**.729**	**.075**	**.268**	**.051**	.042	.047	.048	.724	.159	.683	**.180**	.314	.070	**.877**	.458	.547	**.993**	.433
TA	.131	.310	.083	.531	.036	.122	.040	.051	.065	.091	.909	.376	.853	.171	.307	.057	.807	**.477**	.678	.992	.515
Z_A	.064	.168	.040	.343	.020	.057	.037	.060	.077	.103	.721	.421	.859	.154	.305	.058	.709	.462	.714	.989	.541
$\sqrt{b_1}$	.005	.007	.008	.009	.011	.006	.035	.065	.084	.113	.354	.401	.771	.111	.190	.050	.174	.307	.708	.882	.446
r	.034	.084	.021	.193	.017	.030	.037	.066	.085	.109	.851	.480	.890	.105	.230	.052	.534	.360	.720	.966	.472
R_n	.085	.198	.050	.385	.028	.078	.038	.059	.077	.102	.817	.440	.872	.135	.282	.058	.681	.414	.721	.980	.509
T_{EP}	.073	.149	.045	.267	.034	.065	.042	.058	.071	.090	.807	.417	.866	.129	.284	.062	.580	.368	.722	.952	.488
E_n	.019	.047	.014	.114	.015	.019	.036	.067	.087	.112	.859	.494	.897	.094	.205	.049	.444	.329	.712	.957	.450
M_n	.003	.005	.006	.010	.011	.004	.034	.067	.090	.116	.774	.501	.894	.076	.140	.041	.191	.253	.675	.895	.381
T_{1n}	.018	.029	.017	.046	.020	.019	.040	.057	.070	.089	.261	.292	.645	.152	.301	.058	.448	.436	**.723**	.971	.547
T_{3n}	.074	.212	.043	.409	.037	.070	.039	.061	.083	.110	.775	.482	.896	.098	.205	.045	.646	.333	.695	.971	.433
A^2	.105	.206	.060	.374	.040	.092	.048	.057	.070	.084	.906	.423	.878	.117	.266	.064	.651	.359	.704	.970	.459
SW	.108	.250	.067	.452	.034	.100	.040	.058	.077	.097	.805	.424	.866	.143	.305	.063	.723	.435	.719	.982	.522
SF	.041	.102	.025	.222	.018	.036	.038	.065	.084	.106	.848	.477	.888	.109	.242	.054	.561	.372	.722	.970	.482
SJ	.002	.001	.005	.004	.018	.003	.037	.066	.086	.115	**.930**	**.509**	**.917**	.039	.065	.044	.054	.109	.594	.669	.227
RJB	.002	.002	.004	.003	.011	.003	.036	**.068**	**.092**	**.121**	.819	.507	.902	.065	.119	.041	.091	.206	.666	.784	.348
H_n	.095	.176	.058	.308	.041	.082	.049	.056	.065	.078	.914	.384	.867	.056	**.345**	**.083**	.719	.441	.574	.981	**. 527**

Table 10: *Ranking from first to the fifth of average powers computed from values in Tables 6-7 for* Set 1 *of alternative distributions.*

	Group I Symmetric $(-\infty,\infty)$		Group II Asymmetric $(-\infty,\infty)$		Group III Asymmetric$(0,\infty)$		Group IV (0,1)	
Rank	$n=10$	$n=20$	$n=10$	$n=20$	$n=10$	$n=20$	$n=10$	$n=20$
1	SJ	SJ	H_n	T_{1n}	H_n	Z_A	TV	TV
2	RJB	RJB	T_{1n}	H_n	TV	T_{1n}	TE	TE
3	T_{3n}	M_n	T_EP	Z_A	A	H_n	TV	TA
4	M_n	TZ2	$\sqrt{b_1}$	R_n	T_{1n}	SW	Z_C	QH
5	E_n	E_n	R_n	SW	Z_A	QH	QH	Z_C

Table 11: *Ranking from first to the fifth of average powers computed from values in Tables 8-9 for* Set 2 *of alternative distributions.*

	Symmetric						Asymmetric			
Rank	Short tailed		Close to Normal		Long tailed		Short tailed		Long tailed	
	$n=10$	$n=20$	$n=10$	$n=20$	$n=10$	$n=20$	$n=10$	$n=20$	$n=10$	$n=20$
1	TV	TV	M_n	RJB	SJ	SJ	H_n	H_n	T_{1n}	T_{1n}
2	TA	TA	SJ	M_n	RJB	RJB	TA	TV	SW	SW
3	SW	R_n	RJB	SJ	A^2	SF	TV	TA	R_n	R_n
4	H_n	SW	SF	SF	SF	A^2	SW	SW	H_n	TA
5	A^2	A^2	SW	T_{3n}	T_{3n}	M_n	R_n	R_n	TA	H_n

Tables 10-11 contain the ranking from first to the fifth of the average powers computed from the values in Tables 6-7 and 8-9, respectively. By average powers we can select the tests that are, on average, most powerful against the alternatives from the given groups.

Power against an alternative distribution has been estimated by the relative frequency of values of the corresponding statistic in the critical region for 10000 simulated samples of size n = 10, 20. The maximum reached power is indicated in bold. For computing the estimated powers of the new test, R software is used. We also use R software for computing Pearson chi-square and Shapiro-Francia tests by the package (nortest), command `pearson.test` and `sf.test`, respectively, and also the package (lawstat), command `sj.test` and `rjb.test` for SJ and Robast Jarque-Bera tests, respectively. For the entropy-based test statistics, powers are taken from Zamanzadeh and Arghami (2012) and Alizadeh and Arghami (2011, 2013). In the case of the test based on $\mathrm{H_n}$, we also consider $h_2(x) := x\log(x) - x + 1$ for comparison with $h_1(x) := \left(\frac{x-1}{x+1}\right)^2$.

Results and recommendations

Based on these comparisons, the following recommendations can be formulated for the application of the evaluated statistics for testing normality in practice.

***Set 1* of alternative distributions** (Tables 6-7 and 10): In Group I, for $n = 10$ and 20, it is seen that the tests based on SJ, RJB, T_{3n}, TZ2, M_n and E_n are the most powerful whereas the tests based on I_n, TV, TC and KL are the least powerful. The difference of powers between KL and the others is substantial. In Group II, for $n = 10$ and 20, it is seen that the tests based on H_n, T_{1n}, T_{EP}, R_n, Z_A and $\sqrt{b_1}$ are the most powerful whereas those based on T_{2n}, TV, TC, Kl and Z_w are the least powerful. In Group III, the most powerful tests for $n = 10$ are those based on H_n, TV, TA and T_{1n}, and for $n = 20$, those based on Z_A, T_{1n}, H_n and SW are the most powerful. On the other hand, the least powerful tests are those based on I_n and Z_w are the least powerful. Finally, in group IV, the results are not in favour of the proposed tests. In this group, for $n = 10$ and 20, the most powerful tests are those based on TV, TE, TA, Z_C, Z_A and r, whereas the tests based on TZ_2, SJ and RJB are the least powerful. The SJ and RJB show very poor sensitivity against symmetric distributions in $[0,1]$ such as Unif, $B(2,2)$ or $B(.5,.5)$. For example, for $n = 20$, in the case of the $[0,1]$-Unif alternative, the SJ test has a power of .002 while even the H_n test has a power of .156. From Tables 6-7 one can see that the proportion of times that the SJ and RJB statistics lie below the 5% point of the null distribution are greater than those of the H_n statistic.

Note that for the proposed test, the maximum power in Group II and III was typically attained by choosing h_1.

From the simulation study implemented for *Set 1* of alternative distributions we can lead to different conclusions from that existing in the literature. New and existing results are reported in Table 12.

Table 12: *Comparison of most powerful tests in Groups I–IV, according to Alizadeh and Arghami (2011, 2013) and Zamanzade and Arghami (2012) with new simulation results.*

Alizadeh and Arghami (2011)	JB	SW	KL[a]or SW	KL
Alizadeh and Arghami (2013)	A^2	SW	TA	TV[b]
Zamanzadeh and Arghami (2012)	TZ2	TZ2 or TD	TZ1, KL or TD	KL or TC
New simulation study	SJ or RJB	$H_n$or T_{1n}	H_n or Z_A	TV or TE

[a] Statistic based on Vasicek's estimator
[b] Statistic using nonparametric distribution of Vasicek's estimato

***Set 2* of alternative distributions** (Tables 8-9 and 11): For symmetric short-tailed distributions, it is seen that the tests based on TV, TA and SW are the most powerful. For symmetric close to normal and symmetric long tailed distributions, RJB, JB and M_n are the most powerful. For asymmetric short tailed distributions, H_n, TV and TA are the

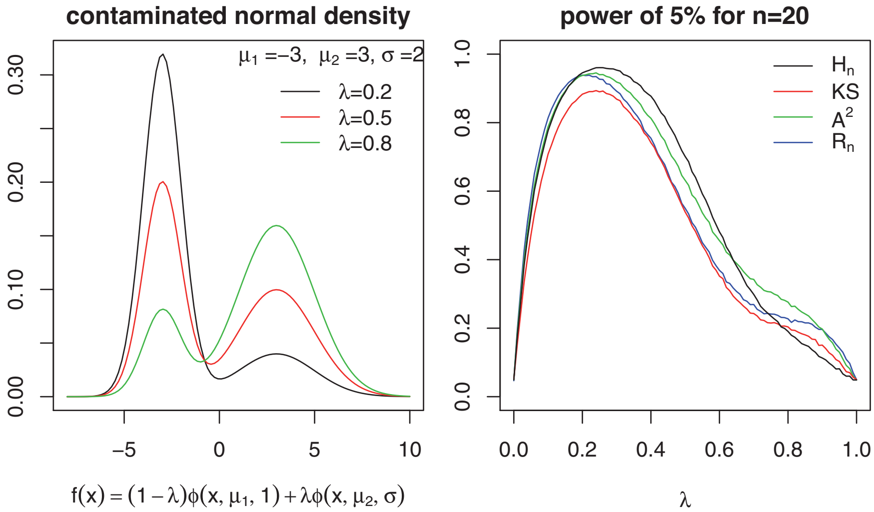

Figure 3: *Left panel: Probability density functions of Contaminated Normal distribution for several values of the parameter* λ*. Right panel: Power of the tests based on* H_n, KS, A^2 *and* R_n *as a function of* λ *against alternative* $\mathrm{CN}(\lambda, \mu_1 = -3, \mu_2 = 3, \sigma = 2)$.

most powerful. Finally, for asymmetric long tailed distributions, T_{1n}, SW and R_n are the most powerful. It is also worth mentioning that the differences between the power of tests based on TV and H_n in $TN(-3,3)$ alternative are not considerable.

In Figure 3 we compare the power of the tests based on H_n, KS, A^2 and R_n against a family of Contaminated Normal alternatives $\mathrm{CN}(\lambda, \mu_1 = -3, \mu_2 = 3, \sigma = 2)$. The left panel of Figure 3 contains the probability density functions of Contaminated Normal alternatives $\mathrm{CN}(\lambda, \mu_1 = -3, \mu_2 = 3, \sigma = 1)$, for $\lambda = .2, .5, .8$, whereas the right panel contains the power comparisons for $n = 20$ and $\alpha = 0.05$. We can see the good power results of H_n for $0.2 < \lambda < 0.6$.

In general, we can conclude that the proposed test H_n has good performance and therefore can be used in practice.

Numerical example

Finally, we illustrate the performance of the new proposal through the analysis of a real data set. One of the most famous tests of normality among practitioners is the Kolmogorov-Smirnov test, mostly because it is available in any statistical software. However, one of its drawbacks is the low power against several alternatives (see also Grané and Fortiana, 2003; Grané, 2012; Grané and Tchirina, 2013).We would like to emphasize this fact through a numerical example.

Armitage and Berry (1987) provided the weights in ounces of 32 newborn babies(see also data set 3 of Henry, 2002, p. 342). The approximate ML estimators of $\hat{\mu} = 111.75$ and $\hat{\sigma} = \sqrt{331.03} = 18.19$. Also sample skewness and kurtosis are $\sqrt{b_1} = -.64$ and

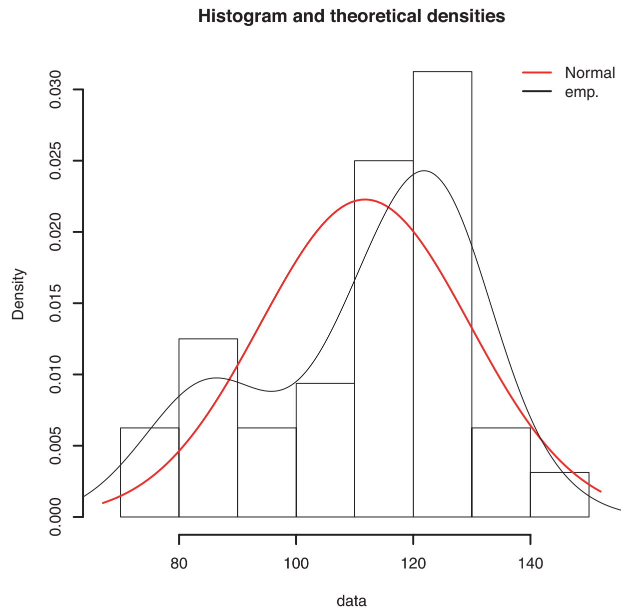

Figure 4: *Histogram and theoretical (normal) distribution for ounces of 32 newborn babies data.*

$b_2 = 2.33$, respectively. From the histogram of these data it can be observed that the birth weights are skewed to the left and may be bimodal (see Figure 4).

When fitting the normal distribution to these data, we find that the KS (Kolmogorov-Smirnov) test does not reject the null hypothesis providing a p-value of 0.093. However with the H_n statistic we are able to reject the null hypothesis of normality at a 5% significance level, since we obtain $\mathrm{H}_n = .0006$ and the corresponding critical value for $n = 32$ is .00047. Also associated p-values of the H_n, SW (Shapiro-Wilk) and SF (Shapiro-Francia) tests are .015, .024 and .036, respectively. Thus, the non-normality is more pronounced by the new test at 5% level. In Appendix, we provide an R software program, to calculate the H_n statistics, the critical points and corresponding p-value.

6. Conclusions

In this paper we propose a statistic to test normality and compare its performance with 40 recent and classical tests for normality and a wide collection of alternative distributions. As expected (Janssen, 2000), the simulation study shows that none of the statistics under evaluation can be considered to be the best one for all the alternative distributions studied. However, the tests based on RJB or SJ have the best performance for symmetric distributions with the support on $(-\infty, \infty)$ and the same happens to TV or TA for distributions with the support on $(0, 1)$. Regarding our proposal, H_n and also T_{1n} are the most powerful for asymmetric distributions with the support on $(-\infty, \infty)$ and distributions with the support on $(0, \infty)$, mainly for small sample sizes.

Acknowledgements

This work has been partially supported by research grant project MTM2014-56535-R (Spanish Ministry of Economy and Competitiveness). The authors are thankful to two Referees and the Editor, whose helpful comments and suggestions contributed to improve the quality of the paper.

Appendix

```
h=function(x) (x-1)^2/(x+1)^2
Hn=function(x) {x=sort(x);n=length(x);
F=pnorm(x, mean(x), sd(x)*sqrt(n/(n-1)))+1;
Fn=1:n/n+1; mean(h(F/Fn))}

##weights in ounces of 32 newborn babies,
data=c(72,80,81,84,86,87,92,94,103,106,107,111,112,115,116,118,
119,122,123,123,114,125,126,126,126,127,118,128,128,132,133,142)
Hn(data)  ## statistics
n=length(data); B=10000; x=matrix(rnorm(n*B, 0, 1), nrow=B, ncol=n)
H0=apply(x, 1, Hn); Q=quantile(H0, .95); Q  ## critical point
length(H0[H0>Hn(data)])/B  ##p-value
```

References

Alizadeh Noughabi, H. and Arghami, N.R. (2012). General treatment of goodness-of-fit tests based on Kullback-Leibler information. *Journal of Statistical Computation and Simulation*, 83, 1–14.

Alizadeh Noughabi, H. and Arghami, N.R. (2010). A new estimator of entropy. *Journal of the Iranian Statistical Society*, 9, 53–64.

Alizadeh Noughabi, H. and Arghami, N.R. (2013). Goodness-of-fit tests based on correcting moments of entropy estimators. *Communications in Statistics-Simulation and Computation*, 42, 499–513.

Alizadeh Noughabi, H. and Arghami, N.R. (2011). Monte carlo comparison of seven normality tests. *Journal of Statistical Computation and Simulation*, 81, 965–972.

Anderson, T.W. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, 49, 765–769.

Anscombe, F.J. and Glynn, W.J. (1983). Distribution of kurtosis statistic b_2 for normal statistics. *Biometrika*, 70, 227–234.

Armitage, P. and Berry, G. (1987). *Statistical Methods in Medical Research*. Blackwell Scientific Publications, Oxford.

Bonett, D.G. and Seier, E. (2002). A test of normality with high uniform power. *Computational Statistics and Data Analysis*, 40, 435–445.

Chen, L. and Shapiro, S.S. (1995). An alternative test for normality based on normalized spacings. *Journal of Statistical Computation and Simulation*, 53, 269–287.

Corea, J.C. (1995). A new estimator of entropy. *Communications in Statistics-Theory and Methods*, 24, 2439–2449.

Cramer, H. (1928). On the composition of elementary errors. *Scandinavian Actuarial Journal*, 11, 141–180.

Csörgo, M. and Révész, P. (1971). *Strong Approximations in Probability and Statistics*. Academic Press, New York.

D'Agostino, R.B. and Stephens, M. (1986). *Goodness-of-fit Techniques*. New York: Marcel Dekker, Inc.

D'Agostino, R.B. and Pearson, E.S. (1973). Tests for departure from normality. Empirical results for the distributions of b_2 and $sqrtb_1$. *Biometrika*, 60, 613–622.

D'Agostino, R.B., Belanger, A. and D'Agostino R.B.Jr. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, 44, 316–321.

D'Agostino, R.B. (1970). Transformation to Normality of the Null Distribution of g_1. *Biometrika*, 57, 679–681.

D'Agostino, R.B. (1971). An omnibus test of normality for moderate and large size samples. *Biometrika*, 58, 341–348.

del Barrio, E., Cuesta-Albertos, J.A., Matrán, C. and Rodríguez-Rodríguez, J.M. (1999). Tests of goodness of fit based on the L2-wasserstein distance. *The Annals of Statistics*, 27, 1230–1239.

de Wet, T. and Venter, J.H. (1972). Asymptotic distributions of certain tests criteria of normality. *South African Statistical Journal*, 6, 135–149.

Dimitriev, Y.G. and Tarasenko, F.P. (1973). On the estimation functions of the probability density and its derivatives. *Theory of Probability & Its Applications*, 18, 628–633.

Doornik, J.A. and Hansen, H. (1994). An omnibus test for univariate and multivariate normality. *Economics Working Papers*, Nuffield College.

Ebrahimi, N., Pflughoeft, K. and Soofi, S.E. (1994). Two measures of sample entropy. *Statistics and Probability Letters*, 20, 225–234.

Epps, T.W. and Pulley, L.B. (1983). A test for normality based on the empirical characteristic function. *Biometrika*, 70, 723–726.

Esteban, M.D., Castellanos, M.E., Morales, D. and Vajda, I. (2001). Monte carlo comparison of four normality tests using different entropy estimates. *Communications in Statistics-Simulation and Computation*, 30, 761–285.

Filliben, J.J. (1975). The probability plot correlation coefficient test for normality. *Technometrics*, 17, 111–117.

Gan, F.F. and Koehler, K.J. (1990). Goodness of fit tests based on P-P probability plots. *Technometrics*, 32, 289–303.

Gel, Y.R., Miao, W. and Gastwirth, J.L. (2007). Robust directed tests of normality against heavy-tailed alternatives. *Computational Statistics and Data Analysis*, 51, 2734–2746.

Gel, Y.R. and Gastwirth, J.L. (2008). A robust modification of the Jarque-Bera test of normality. *Economics Letters*, 99, 30–32.

Grané, A. and Fortiana, J. (2003). Goodness of fit tests based on maximum correlations and their orthogonal decompositions. *Journal of the Royal Statistical Society. Series B: Methodological*, 65, 115–126.

Grané, A. (2012). Exact goodness of fit tests for censored data. *Annals of the Institute of Statistical Mathematics*, 64, 1187–1203.

Grané, A. and Tchirina, A. (2013). Asymptotic properties of a goodness-of-fit test based on maximum correlations. *Statistics*, 47, 202–215.

Henry, C. T. (2002). *Testing for Normality*. Marcel Dekker, New York.

Janssen, A. (2000). Global power functions of godness-of-fit tests. *Annals of Statististics*, 28, 239–253.

Jarque, C.M. and Bera, A.K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6, 255–259.

Jarque, C.M. and Bera, A.K. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review/Revue Internationale de Statistique*, 55, 163–172.

Kolmogorov, A.N. (1933). Sulla determinazione empirica di una legge di dislibuzione. *Giornale dell'Istitu Italiano degli Attuari*, 4, 83–91.

Krauczi, E. (2009). A study of the quantile correlation test for normality. *TEST*,18, 156–165.

Kuiper, N.H. (1962). Test concerning random points on a circle. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen Series A*, 63, 38–47.

LaRiccia, V. (1986). Optimal goodness-of-fit tests for normality against skewness and kurtosis alternatives. *Journal of Statistical Planning and Inference*, 13, 67–79.

Lilliefors, H.W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62, 399–402.

Louter, A.S. and Koerts, J. (1970). On the Kuiper test for normality with mean and variance unknown. *Statistica Neerlandica*, 24, 83–87.

Martinez, J.and Iglewicz, B. (1981). A test for departure from normality based on a biweight estimator of scale. *Biometrika*, 68, 331–333.

Mises, R. von (1931). *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik*. Deuticke, Leipzig and Vienna.

Pardo, L. (2006). *Statistical Inference based on Divergence Measures*. New York: Taylor Francis Group.

Park, S. and Park, D. (2003). Correcting moments for goodness-of-fit tests based on two entropy estimates. *Journal of Statistical Computation and Simulation*, 73, 685–694.

Pettitt, A.N. (1977). A Cramer-von Mises type goodness of fit statistic related to $\sqrt{b_1}$ and b_2. *Journal of the Royal Statistical Society. Series B: Methodological*, 39, 364–370.

Poitras, G. (2006). More on the correct use of omnibus tests for normality. *Economics Letters*, 90, 304–309.

Romao, X., Delgado, R. and Costa, A. (2010). An empirical power comparison of univariate goodness-of-fit tests for normality. *Journal of Statistical Computation and Simulation*, 80, 545–591.

Shao, J. (2003). *Mathematical Statistics*. New York: Springer, Verlag.

Shapiro, S.S., Wilk, M.B. and Chen, M.H.J. (1968). A comparative study of various tests for normality. *Journal of the American Statistical Association*, 63, 1343–1372.

Shapiro, S.S. and Wilk, M.B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52, 591–611.

Shapiro, S.S. and Francia, R.S. (1972). An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, 67, 215–216.

Shenton, L. and Bowman, K. (1977). A bivariate model for the distribution of b_1 and b_2. *Journal of the American Statistical Association*, 72, 206–211.

Van Es, B. (1992). Estimating functionals related to a density by class of statistics based on spacing. *Scandinavian Journal of Statistics*, 19, 61–72.

Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of the Royal Statistical Society. Series B: Methodological*, 38, 54–59.

Watson, G.S. (1961). Goodness-of-fit tests on a circle. *Technometrics*, 48, 109–114.

Yap, B.W. and Sim, C.H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, 81, 2141–2155.

Yazici, B. and Yolacan, S. (2007). A comparision of various tets of normality. *Journal of Statistical Computation and Simulation*, 77, 175–183.

Zamanzade, E. and Arghami, N.R. (2012). Testing normality based on new entropy estimators. *Journal of Statistical Computation and Simulation*, 82, 1701–1713.

Zhang, J. and Wu, Y. (2005). Likelihood-ratio tests for normality. *Computational Statistics and Data Analysis*, 49, 709–721.

Zhao, J. and Xu, X. (2014). Goodness-of-fit tests for location scale families based on random distance. *Journal of Statistical Computation and Simulation*, 84, 739–752.